

47-13582

## \* NOTICES \*

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So, the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

## CLAIMS

## [Claim(s)]

[Claim 1] The portable speech-recognition output auxiliary device carry out having voice-input/output equipment which attached the voice-input means of the shape of flatness which detects vibration which generates from the oscillating generating object concerned in the rear-face side of the wearing object of the shape of a strip of paper formed of the material of the absorption-of-sound nature which twists and fixes to an oscillating generating object, and changes into an electric oscillation-frequency signal, and attached the voice output means of the shape of flatness which outputs the sound signal according to the aforementioned oscillation-frequency signal to the front-face side of the aforementioned wearing object further as

[Claim 2] The portable speech recognition output auxiliary device characterized by providing the following. Voice-input/output equipment which has a voice output means to output the sound signal according to the oscillation frequency signal inputted by voice input means to detect vibration generated from an oscillating generating object, and to output an electric oscillation frequency signal, and this voice input means. The speech recognition section which recognizes a voiceprint, the strength of sound, height, and generating sound from the oscillation frequency signal inputted from the aforementioned voice input means. When the aforementioned voice sign which the voice sign corresponding to two or more standard voice patterns and patterns concerned is memorized beforehand, and has already been remembered to be a voice sign concerning the generating sound recognized by the aforementioned speech recognition section is compared and both the voice sign is in agreement. The voice sign judging section which outputs the speech information which reads the standard voice pattern corresponding to the aforementioned voice sign, and consists of strength, height, etc. of the standard voice pattern concerned, the aforementioned voiceprint, and sound.

[Claim 3] The portable speech-recognition output auxiliary device which carries out [ having added the speech-synthesis section which compounds the aforementioned standard voice pattern and the aforementioned voiceprint outputted from the aforementioned voice sign judging section in claim 2 publication, attaches both the strength of the aforementioned sound, and both / either or / further, and creates composite tone, and the voice conversion output section which change into a sound signal the composite tone created in this speech-synthesis section, and output from the aforementioned voice output means, and ] as the feature.

[Claim 4] The speech synthesis section which compounds the aforementioned standard voice pattern and the aforementioned voiceprint which are outputted from the aforementioned voice sign judging section in claim 2 publication, attaches the strength and height of the aforementioned sound further, and creates composite tone. The phonetic-memory section which memorizes the composite tone created by this speech synthesis section. The voice conversion output section which changes into a sound signal the composite tone memorized by this phonetic-memory section, and is outputted from the aforementioned voice output means. The voice recurrence switch which reads the composite tone memorized by the aforementioned phonetic-memory section, and is made to output repeatedly from the aforementioned voice output means. The portable speech recognition output auxiliary device characterized by adding the voice adjustable means which carries out adjustable [ of both the speed of the sound signal outputted from the aforementioned voice conversion output section, and both / either or ].

[Claim 5] The speech recognition section which recognizes the strength of a voiceprint and sound and height, and generating sound from the oscillation frequency signal inputted as the voice-input/output equipment portion which has a voice input means and a voice output means from the aforementioned voice input means. Two or more voice signs beforehand remembered to be the voice signs of the generating sound recognized by

this speech recognition section are compared. When both the voice sign is in agreement, the standard voice pattern corresponding to the voice sign concerned memorized beforehand is read. The voice sign judging section which generates speech information, such as strength of this standard voice pattern, the aforementioned voiceprint, and sound, and height A main frame portion with the speech synthesis section which compounds strength, height, etc. of these standard voice pattern, a voiceprint, and sound, and the voice conversion output section which changes into a sound signal the composite tone created by this speech synthesis section, and is outputted from the aforementioned voice output means. The voice repeat switch which reads the composite tone memorized by the aforementioned phonetic-memory section, and is made to output repeatedly from the aforementioned voice output means. The portable speech recognition output auxiliary device characterized by dividing into a part for a voice controller with the voice adjustable means which carries out adjustable [ of both the speed of the sound signal outputted from the aforementioned voice conversion output section, and both / either or ].

---

[Translation done.]

**\* NOTICES \***

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

**DETAILED DESCRIPTION**

---

**[Detailed Description of the Invention]****[0001]**

[Industrial Application] this invention relates to the portable speech recognition output auxiliary device which those who solve and utter not clear voice, such as for example, a vocal cord extraction person and a physically handicapped person, use, and makes carrying possible especially easily with respect to a suitable portable speech recognition output auxiliary device, and recognizes not clear voice appropriately, and can be used for assistance of conversation.

**[0002]**

[Description of the Prior Art] The conventional voice recognition unit carries out comparison collating of the standard voice pattern of a large number which memorize the standard voice pattern of a large number corresponding to the voice which a healthy person utters beforehand, and are beforehand remembered to be the voice patterns uttered from the healthy person's mouth, and if there is a standard voice pattern which is in agreement with the voice pattern which a healthy person utters, recognizing the voice which a healthy person utters from the standard voice pattern concerned is performed.

[0003] There are a method reproduced when it is equipment with which a voice synthesizer outputs voice on the other hand, and recording the voice which announcer utters, compressing and recording it on a low bit by tools of analysis and outputting further, and a rule composite system which combines a single sound corresponding to the Japanese syllabary to input, and superimposes an accent and intonation. The former is used as an output of an audio response unit, and is used in the order entry field combining PB input of a push-button phone. The technology changed into direct-sound voice from the text of Japanese and English is developed, and the place of the latter expected to development of future technology is large.

**[0004]**

[Problem(s) to be Solved by the Invention] However, the conventional voice recognition unit cannot be recognized at all about the voice uttered by those who are performing speech recognition from the voice pattern uttered from a healthy person's mouth, for example, extracted vocal cords by the operation etc., those who lost the tongue by the lingual cancer, the non-healthy person who utters not clear voice. Since the reason does not utter voice from a mouth from the first in the case of those who extracted vocal cords since recognition not only becomes impossible, but air vibration of the voice uttered from a mouth was detected in order to utter not clear voice, or those who lost the tongue by the lingual cancer, it becomes impossible applying it.

[0005] In addition, although many and unspecified speech recognition will become possible depending on a future technical progress situation and the various equipments which used the voice recognition unit will be used in everyday life, it is thought anyway that it is development of equipment effective in a healthy person. Therefore, since an audio generating speed is slow, even if the voice of the non-healthy person with various handicaps is not clear, or new equipment is developed with much trouble, it is thought that it is very difficult to fully master it.

[0006] On the other hand, in the case of the aforementioned voice synthesizer, if the sound signal which was filled with much language which an individual utters, or feeling does not become but it sees from a viewpoint of talking, it is still inadequate.

[0007] this invention was not made in view of the above-mentioned actual condition, and aims at those who cannot utter voice from a mouth offering the portable speech recognition output auxiliary device which can be inputted certainly for the signal equivalent to voice.

[0008] Moreover, other purposes of this invention recognize correctly the not clear voice which a non-healthy person utters, and are to offer the portable speech recognition output auxiliary device which realizes speech synthesis including the feeling of those who utter voice.

[0009] Furthermore, other purposes of this invention are to offer the portable speech recognition output auxiliary device which generates a suitable sound signal, taking into consideration the situation of a non-healthy person's body. Furthermore, other purposes of this invention are to offer the portable speech recognition output auxiliary device which the non-healthy person could equip easily and was rich in operability.

[0010]

[Means for Solving the Problem] In order to solve the above-mentioned technical problem, invention corresponding to a claim 1 The voice input means of the shape of flatness which detects vibration generated from the oscillating generating object concerned in the rear-face side of the wearing object of the shape of a strip of paper formed with the cloth of the absorption-of-sound nature twisted and fixed to an oscillating generating object, and is changed into an electric oscillation frequency signal is attached. It is the portable speech recognition output auxiliary device which has voice-input/output equipment which attached the voice output means of the shape of flatness which furthermore outputs the sound signal according to the aforementioned oscillation frequency signal to the front-face side of the aforementioned wearing object.

[0011] Next, the voice-input/output equipment which has a voice output means to output the sound signal according to the oscillation frequency signal inputted by voice input means for invention corresponding to a claim 2 to detect vibration generated from an oscillating generating object, and to output an electric oscillation frequency signal, and this voice input means, The speech recognition section which recognizes the strength of a voiceprint and sound and height, and generating sound from the oscillation frequency signal inputted from the aforementioned voice input means, When the aforementioned voice sign which the voice sign corresponding to two or more standard voice patterns and patterns concerned is memorized beforehand, and has already been remembered to be a voice sign concerning the generating sound recognized by the aforementioned speech recognition section is compared and both the voice sign is in agreement, It is the portable speech recognition output auxiliary device which prepared the voice sign judging section which outputs the speech information which reads the standard voice pattern corresponding to the aforementioned voice sign, and consists of strength, height, etc. of the standard voice pattern concerned, the aforementioned voiceprint, and sound.

[0012] Invention corresponding to a claim 3 to next, the requirements for composition of invention corresponding to a claim 2 The speech synthesis section which compounds the aforementioned standard voice pattern and the aforementioned voiceprint which are newly outputted from the aforementioned voice sign judging section, attaches the strength and height of the aforementioned sound further, and creates composite tone, It is the portable speech recognition output auxiliary device which comes to add the voice conversion output section which changes into a sound signal the composite tone created in this speech synthesis section, and is outputted from the aforementioned voice output means.

[0013] Invention corresponding to a claim 4 to furthermore, the requirements for composition of invention corresponding to a claim 2 The speech synthesis section which compounds the aforementioned standard voice pattern and the aforementioned voiceprint which are newly outputted from the aforementioned voice sign judging section, attaches the strength and height of the aforementioned sound further, and creates composite tone, The voice storage section which memorizes the composite tone created by this speech synthesis section, The voice conversion output section which changes into a sound signal the composite tone memorized by this voice storage section, and is outputted from the aforementioned voice output means, The voice recurrence switch which reads the composite tone memorized by the aforementioned voice storage section, and is made to output repeatedly from the aforementioned voice output means, It is the portable speech recognition output auxiliary device which comes to add the voice speed adjustable means which carries out adjustable [ of the speed of the sound signal outputted from the aforementioned voice conversion output section ], and the voice strength adjustable means which carries out adjustable [ of the sound signal level outputted from the aforementioned voice conversion output section ], and attaches strength.

[0014] Furthermore, the voice-input/output equipment portion into which invention corresponding to a claim 5 has a voice input means and a voice output means, A main frame portion with the speech recognition section, the voice sign judging section, and the voice conversion output section, The voice repeat switch which reads the composite tone memorized by the aforementioned voice storage section, and is made to

output repeatedly from the aforementioned voice output means. It is the portable speech recognition output auxiliary device divided into a part for a voice controller with the voice speed adjustable means which carries out adjustable [ of the speed of the sound signal outputted from the aforementioned voice conversion output section ], and the voice strength adjustable means which carries out adjustable [ of the sound signal level outputted from the aforementioned voice conversion output section ], and attaches strength.

[0015]

[Function] Invention corresponding to a claim 1 therefore, by having provided the above meanses By having used the cloth of absorption-of-sound nature for the oscillating generating object, for example, the wearing object twisted and fixed to a non-healthy person's neck, and having attached flatness-like a voice input means and a voice output means in the rear-face [ of a wearing object ], and front-face side individually, respectively The influence of noise which enters from the voice uttered from a mouth or the outside can be prevented, moreover the burden to a non-healthy person's throat is mitigated, and vibration uttered from a direct throat can be inputted certainly.

[0016] Next, the speech recognition section recognizes a voiceprint, the strength of sound, the height of sound, and utterance sound from the oscillation frequency signal inputted from a voice input means, and sends out invention corresponding to a claim 2 to the voice sign judging section. In this voice sign judging section, since the voice sign corresponding to two or more standard voice patterns and patterns concerned is memorized beforehand When the voice sign already remembered to be a voice sign concerning the generating sound sent from the speech recognition section is compared and both the voice sign is in agreement, Since the speech information which reads the standard voice pattern corresponding to the voice sign, and consists of strength, height, etc. of the standard voice pattern concerned, the aforementioned voiceprint, and sound is outputted The standard voice pattern changed into the long language used for conversation etc. every day from the short language which can recognize correctly also with the not clear voice which a non-healthy person utters, and a non-healthy person utters can be outputted easily.

[0017] Furthermore, invention corresponding to a claim 3 has the same operation as invention corresponding to a claim 2, and also Since the standard voice pattern sent from the voice sign judging section in the speech synthesis section and the aforementioned voiceprint are compounded, the strength of sound and height are attached further and composite tone is created Since it can synthesize voice including feeling, composite tone is moreover changed into a sound signal in the sound signal conversion output section and it outputs from the aforementioned voice output means, the sound signal accompanied by feeling expression can be outputted.

[0018] Furthermore, invention corresponding to a claim 4 can output the same sound signal, without uttering voice from the beginning, even when asked again by the partner, since it has the same operation as invention corresponding to a claim 2 and a claim 3, and also a voice recurrence switch is operated, composite tone is again read from the aforementioned voice storage section and it outputs repeatedly from a voice output means. Moreover, a sound signal can be outputted at a speed intelligible for a healthy person by carrying out adjustable [ of the output speed of a sound signal ] by the voice speed adjustable means. Moreover, since adjustable [ of the sound signal level ] is carried out and strength is attached and outputted by the voice strength adjustable means, a sound signal intelligible for a healthy person can be outputted similarly.

[0019] Furthermore, the voice-input/output equipment portion into which invention corresponding to a claim 5 has a voice input means and a voice output means. By dividing into a part for a main frame portion with the speech recognition section, the voice sign judging section, the sound signal conversion output section, etc., and a voice controller with various adjustment functions A voice-input/output equipment portion is twisted around a non-healthy person's neck, the pendant of the main frame portion is carried out to a part for the waist of a fuselage etc., it can carry easily and a part for a voice controller can be easily operated, if it has at hand and is made to operate it.

[0020]

[Example] Hereafter, the example of this invention is explained with reference to a drawing. Drawing 1 is the block diagram showing the composition of this invention equipment. In \*\*\*\*, 1 is voice-input/output equipment, and as this shows drawing 2, a wearing object 11 of cloth like the corset twisted around a neck is used at the times, such as a whiplash. A voice output means 13 to output the voice input means 12 and sound signal which incorporate directly vibration uttered from a throat is attached in the proper part of this wearing object 11, and since it twists and fixes to a neck further, pieces of Velcro 14a and 14b are attached in the confrontation side of wearing object both ends. In addition, you may fix using various fixed meanses, for

example, hook etc., of the conventional common knowledge of those other than piece-of-Velcro 14a and 14b etc.

[0021] The aforementioned wearing object 11 is created with the cloth which was excellent in absorption-of-sound nature, and absorbs the noise which enters from the voice uttered from a mouth by this, or the outside, and it is made not to affect the aforementioned voice input means 12 like the curtain ground which intercepts external noise. The voice input means 12 is evenly attached in the background (inside) side section of the wearing object 11, and changes and outputs vibration uttered from a throat to an electrical signal. Thus, by carrying out flattening, it is easy to get used to the wearing object 11, and there is no feeling of oppression to a throat as a result the burden to a throat can be mitigated. On the other hand, a flat loudspeaker is attached by the voice output means 13 like [ the voice input means 12 ] the side-front (outside) side section of an opposite side 11, i.e., a wearing object. Thus, by attaching a flat loudspeaker in the transverse plane on the same vertical line as a mouth, the burden to a throat is mitigated, and if it talks and sees from a partner, the state where voice emits from a mouth will be made. Moreover, this voice output means 13 shall be attached in the natural installation state which is not conspicuous as long as it is all out by devising a proper device to a wrap and the color of the voice output means 13 for the wearing object 11, an affiliated color, or a proper material.

[0022] 2 is the speech recognition section which recognizes the feature of an individual voiceprint, the strength of sound and height, and the right utterance from the voice oscillation frequency signal inputted from the voice input means 12. This speech recognition section 2 consists of the voice spectrum conversion means 21, the tone-quality judging means 22, a voiceprint judging means 23, and utterance sound recognition means 24 grade, as shown in drawing 3. This voice spectrum conversion means 21 is changed into a voice spectrum as shown in drawing 4 (b) by sampling a voice oscillation frequency signal as shown in drawing 4 (a) with a predetermined period. Although it means in [ each / of a voice spectrum / component / the vertical direction from reference level ] level how much the tone-quality judging means 22 judges the strength of a voice spectrum to sound, and height, and reference level beforehand predetermined in the strength of sound was set up, among those it is separated and the height of sound is dependent on the frequency of sound on the other hand, the level of each [ of a voice spectrum ] component is expressed chiefly here. The voiceprint judging means 23 extracts the frequency component level of a voice spectrum, and the utterance sound recognition means 24 determines utterance sound from the distribution state of a voice spectrum, and it changes and outputs it to the character code corresponding to the utterance sound, for example, "A", "I", and that code. And the data judged by these judgment meanses 22-24 are outputted serially, and are sent to the voice sign judging section 3.

[0023] This voice sign judging section 3 compares the voice sign already remembered to be this voice sign, both the voice sign becomes the same, solves [ a standard voice pattern and the voice sign corresponding to it are memorized beforehand, the character code (voice sign) which is the right utterance sound by which speech recognition was carried out with the utterance sound recognition means 24 is taken out, ] it, and it has the function which outputs the standard voice pattern corresponding to it. It is constituted by a voice pattern storage means 31 to specifically memorize a standard voice pattern as shown in drawing 5, a voice sign storage means 32 to memorize the voice sign corresponding to each standard voice pattern of this voice pattern storage means 31, and the voice sign judging means 33.

[0024] This voice sign judging means 33 changes the data about the strength of the sound from the aforementioned tone-quality judging means 22, and height, and the feature data of the voiceprint from the voiceprint judging means 23 into the state of the waiting for a buffer memory. About the voice sign of the right utterance sound recognized with the utterance sound recognition means 24 if the same as that of the voice sign which carries out the comparison reference of the voice sign of a large number memorized by the voice sign concerned and the voice sign storage means 32, and has already been memorized The standard voice pattern corresponding to a voice sign is taken out from the voice pattern storage means 31, and it memorizes in the speech information storage section 4 with the data which are already in the state of the waiting for a buffer memory. At this time, you may also memorize simultaneously the voice sign of the generating sound of the generating sound recognition means 24. When the voice sign already remembered to be the voice sign recognized by the utterance sound recognition means 24 on the other hand becomes inharmonious, the voice sign of the utterance sound recognized with the utterance sound recognition means 24 is outputted.

[0025] In addition, the standard voice pattern memorized by the aforementioned voice pattern storage means

31 is a pattern equivalent to the language used in everyday conversation, such as "good morning", "thank you", and "good-bye." That is, by changing into long language from a short voice sign, even if a non-healthy person does not utter all language, it is patternizing possible [ conversation ] enough.

[0026] After the aforementioned speech information storage section 4 stores temporarily speech information, such as a voice sign of the generating sound recognized if needed [ a standard voice pattern and if needed ] concerning the feature of a voiceprint, the strength of sound, the height of sound, and utterance sound, it is sent out to the speech synthesis section 5.

[0027] A speech information storage means 51 to memorize the speech information sent from the speech information storage section 4 in this speech synthesis section 5 as shown in drawing 6. By compounding a standard voice pattern and the feature data of a voiceprint among the speech information memorized by this speech information storage means 51, and attaching the strength of sound, and the height of sound to the composite tone which starts further. The composite tone recovery-ized completely is made and it consists of speech synthesis meanses 52 to memorize in the consecutive voice storage section 6.

[0028] 7 is the voice conversion output section and this has the function which reads the composite tone information memorized by the voice storage section 6, changes into the analog signal in which a voice output is possible, and outputs voice from the voice output means 13.

[0029] Furthermore, the voice output controller 8 is formed in this equipment. The reason for having formed this voice output controller 8 has the contents of conversation in enabling it to transmit suitable for the other party according to a non-healthy person's situation. That is, when the sound signal outputted from the voice output means 13 is asked again from the other party at once, in order to make the voice output controller 8 output composite tone repeatedly from the voice storage section 6, the voice recurrence switch 81 which performs read-out operation is formed in it. Since it is very serious that a non-healthy person utters the same voice from the beginning, this is for mitigating the burden.

[0030] Moreover, the voice speed adjustable machine 82 and the voice strength adjustable machine 83 are formed in this voice output controller 8. It carries out adjustable [ of the speed of a sound signal ] by incorporating the analog primary phase lead lag network which used the capacitor etc., and connecting a primary phase lead lag network with the voice conversion output section 7 side too hastily suitably with the voice speed adjustable vessel 82 beforehand. Since a non-healthy person's utterance speed does not necessarily have early, this is for making it the speed which changes suitably the output speed of the composite tone outputted from the voice output means 13, and a healthy person tends to catch. Moreover, the voice strength adjustable machine 83 attaches and outputs strength to a sound signal by carrying out adjustable [ of the level of the sound signal by the side of the voice conversion output section 7 ], or carrying out adjustable [ of the amplification factor ]. This is for making it easy to attach strength to the sound signal outputted from the voice output means 13 even place [ with much external noise ], and to catch.

[0031] Next, operation of the equipment constituted as mentioned above is explained. First, after a non-healthy person twists the wearing object 11 of voice-input/output equipment 1 around a neck, the piece-of-Velcro portion prepared in the ends confrontation side of the wearing object 11 is forced, and it fixes. At this time, it sets up so that the voice output means 13 attached in the wearing object 11 may come to a transverse-plane position, and the voice input means 12 is set as the position of the flank of the part which is the easiest to incorporate vibration of a throat, for example, a neck. Since the voice input means 12 and the output means 13 are formed in the shape of flatness at this time, it is easy to get used to a neck, and the burden to a throat decreases very much.

[0032] If a non-healthy person generates voice in this state, vibration of the throat of the non-healthy person concerned is incorporated with the voice input means 12, and it changes into an electric oscillation frequency signal, and sends out to the speech recognition section 2.

[0033] Here, after the speech recognition section 2 changes into a voice spectrum the oscillation frequency signal inputted from the voice input means 12 by the voice spectrum conversion means 21, it is sent out to the tone-quality judging means 22, the voiceprint judging means 23, and the generating sound judging means 24. Each [ these ] judgment meanses 22-24 determine the feature and the right generating sound of the strength of sound and the height of sound, and a voiceprint according to the criteria mentioned above, and especially, in the case of generating sound, they are changed at the character code (voice sign) corresponding to generating sound, and they send them out to the voice sign judging section 3 with the strength of sound and the height of sound, and the feature data of a voiceprint.

[0034] In this sign judging section 3, a standard voice pattern is beforehand memorized by the voice pattern

storage means 31, and the voice sign corresponding to the aforementioned standard voice pattern is memorized by the voice sign storage means 32, and it is especially saved to the standard voice pattern in the form of the pattern equivalent to the language used in everyday conversation, such as "good morning", "thank you", and "good-bye."

[0035] Therefore, when the character code (voice sign) which is the right phonation sound recognized by the speech recognition section 2 in the sign judging section 3 was received, the voice sign already remembered to be some of the voice sign is compared and both the voice sign becomes the same. The standard voice pattern corresponding to it is read, and it sends out to the speech synthesis section 5 through the speech information storage section 4 with the data about the strength of the sound from the aforementioned tone-quality judging means 22, and height, and the feature data of the voiceprint from the voiceprint judging means 23.

[0036] Here, once the speech synthesis section 5 memorizes speech information, such as a standard voice pattern sent from the speech information storage section 4, strength of sound, height, and a voiceprint, for the speech information storage means 51, it synthesizes voice with the speech synthesis means 52. After this speech synthesis compounds a standard voice pattern and the feature data of a voiceprint among speech information, makes the composite tone which attached the strength of sound, and the height of sound to the composite tone which starts further, and made the perfect recovery and memorizes it in the phonetic-memory section 6, it is sent to the voice conversion output section 7. In this voice conversion output section 7, the composite tone information memorized by the phonetic-memory section 6 is read, it changes into the analog signal in which a voice output is possible, and voice is outputted from the voice output means 13.

[0037] When asked again at this time, for example, the other party, if the voice recurrence switch 81 is operated, since a non-healthy person will read composite tone information from the phonetic-memory section 6 again, will change into the analog signal in which a voice output is possible in the voice conversion output section 7 and will output voice from the voice output means 13, he can tell the suitable sound signal for the other party, i.e., the content of conversation. Moreover, if output speed of a sound signal is suitably carried out early with the voice speed adjustable vessel 82 when a non-healthy person's phonation speed is slow, it will become easy to catch a healthy person etc. Moreover, sound signal level will be enlarged, and it can output from the voice output means 13, and will become easy to catch a healthy person etc., if adjustable operation of the voice strength adjustable machine 83 is carried out, when there is much external noise similarly, for example.

[0038] Therefore, since the wearing object 11 which should turn into a main part of voice-input/output equipment 1 was created with the cloth excellent in absorption-of-sound nature etc. according to the composition of the above examples, when it twists around a non-healthy person's neck, the noise which it not only gets used completely, but enters from the voice uttered from a mouth or the outside is absorbed, and vibration uttered from a throat can be appropriately inputted from the voice input means 12. And if it attaches so that flatness-like the voice input means 12 and the voice output means 13 may be stuck on the field section of the wearing object 11, it is convenient to carry at lightweight C, the feeling of oppression over a throat etc. is lost, and the burden to a throat can be mitigated. Moreover, the feature of a voiceprint, the strength of sound and the height of sound, and utterance sound are recognized from the oscillation frequency signal inputted from the voice input means 12 in the speech recognition section 2. When the voice sign of this utterance sound, the feature of a voiceprint, the strength of sound, the height information on sound, etc. are sent out to the voice sign judging section 3, the voice sign of a large number beforehand remembered to be voice signs here is compared and both the voice sign is in agreement. The standard voice pattern equivalent to the language used in everyday conversation, such as "thank you" corresponding to the voice sign concerned and "good-bye", is read. Since it was made to send out to the speech synthesis section 5 with the feature of the aforementioned voiceprint, the strength of sound, the height of sound, etc., the standard voice pattern of the long letter which is conversation every day can be outputted from utterance of the first short conversation by the non-healthy person, and the burden of the conversation by the non-healthy person can fully be assisted.

[0039] Furthermore, in the speech synthesis section 5, since the feature of a voiceprint was compounded to the standard voice pattern out of the speech information and the strength of sound and the height of sound were further attached once memorizing various kinds of speech information sent from the voice sign judging section 3 side, composite tone including a non-healthy person's sentiment can be created.

[0040] Furthermore, a suitable sound signal can be outputted, operating it suitably according to a non-healthy



person's situation, or the listening comprehension state of the other party, since the voice recurrence switch 81 which repeats and outputs a sound signal, and the voice speed adjustable machine 82 and the voice strength adjustable machine 83 which carries out adjustable [ of the speed or intensity of a sound signal ] were formed.

[0041] In addition, although the above-mentioned example described the whole composition, when it thinks from a viewpoint which a non-healthy person carries easily and operates simply, considering as the following division composition is desirable. That is, the voice-input/output equipment portion which has the voice input means 12 and the voice output means 13, A part for the equipment book soma containing a part for the power supply section which consists of the speech recognition section 2, the voice sign judging section 3, the speech information storage section 4, the speech synthesis section 5, the phonetic-memory section 6, and voice conversion output section 7 grade. If it divides into a part for the voice output controller of the voice recurrence switch 81, the voice speed adjustable machine 82, and voice strength adjustable machine 83 grade. If it is made to connect by the signal line suitably, a voice-input/output equipment portion is twisted around a neck, it fixes, the pendant of the part for an equipment book soma is carried out to the waist, and, as for a hamper, it is made by the amount of voice output controller in a hand, and by this, it can carry easily and operability can be raised.

[0042] Moreover, although cloth was used, the wearing object 11 will not be especially limited, if it is a material similar to \*\*\*\* of absorption-of-sound nature, or it. In addition, in the range which does not deviate from the summary, this invention deforms variously and can be carried out.

[0043]

[Effect of the Invention] As explained above, according to this invention, the following various effects are done so. In invention of a claim 1, vibration of a throat can be inputted appropriately, without those who cannot utter voice from a mouth also being able to input the signal equivalent to voice certainly, and pressing a non-healthy person's throat.

[0044] Invention of claims 2 and 3 can recognize correctly the not clear voice which a non-healthy person utters, and can realize speech synthesis including the feeling of those who utter voice by moreover compounding the strength of a voice pattern, a voiceprint, and sound etc.

[0045] Next, invention of a claim 4 can perform voice operation suitably in consideration of the situation of a non-healthy person's body according to a partner's listening comprehension condition, and a proper sound signal can be generated. Furthermore, by dividing composition appropriately, a non-healthy person can equip with invention of a claim 5 easily, and it can raise the operability by the non-healthy person.

---

[Translation done.]

\* NOTICES \*

Japan Patent Office is not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

DESCRIPTION OF DRAWINGS

---

[Brief Description of the Drawings]

[Drawing 1] The functional block diagram showing one example of the portable speech recognition output auxiliary device concerning this invention.

[Drawing 2] Drawing showing the composition of the voice-input/output equipment shown in drawing 1.

[Drawing 3] The functional block diagram which materialized the speech recognition section shown in drawing 1.

[Drawing 4] Drawing explaining the speech recognition by the speech recognition section.

[Drawing 5] The functional block diagram which materialized the voice sign judging section shown in drawing 1.

[Drawing 6] The functional block diagram which materialized the speech synthesis section shown in drawing 1.

[Description of Notations]

1 [ — The voice sign judging section, 4 / -- The speech information storage section, 5 / — The speech synthesis section, 6 / -- The phonetic-memory section, 7 / — The voice conversion output section, 8 / — A voice output controller, 11 / — A wearing object, 12 / -- A voice input means, 13 / -- A voice output means, 81 / -- A voice recurrence switch, 82 / -- A voice speed adjustable machine, 83 / -- Voice strength adjustable machine. ] — Voice-input/output equipment, 2 —

---

[Translation done.]



(2)

特開平7-13582

2

【発明の概要】

【請求項1】 振動発生体に巻付け固定する吸音性の素材によって形成された筒状の装着体の裏面側に当該振動発生体から発生する振動を検出して電気的な振動周波数信号に変換する平坦状の音声入力手段を取り付け、さらに前記装着体の表面側に前記振動周波数信号に応じた音声信号を出力する平坦状の音声出力手段を取り付けた音声入出力装置を有することを特徴とする携帯用音声認識出力補助装置。

【請求項2】 振動発生体から発生する振動を検出して電気的な振動周波数信号を出力する音声入力手段およびこの音声入力手段によって入力された振動周波数信号に応じた音声信号を出力する音声出力手段とを有する音声入出力装置と、前記音声入力手段から入力された振動周波数信号から声紋、音の強弱、高低、発生音を認識する音声認識部と、予め複数の標準音声パターンおよび当該パターンに対応する音声符号が記憶され、前記音声認識部によって認識された発生音に係わる音声符号と既に記憶されている前記音声符号とを比較し、両音声符号が一致したとき、前記音声符号に対応する標準音声パターンを読み出し、当該標準音声パターン、前記声紋、音の強弱および高低等からなる音声情報を出力する音声符号判定部とを備えたことを特徴とする携帯用音声認識出力補助装置。

【請求項3】 請求項2記載において、前記音声符号判定部から出力される前記標準音声パターンと前記声紋とを合成し、さらに前記音の強弱および高低の何れか一方または両方を付して合成音を作成する音声合成部と、この音声合成部で作成された合成音を音声信号に変換して前記音声出力手段から出力する音声変換出力部とを付加したことを特徴とする携帯用音声認識出力補助装置。

【請求項4】 請求項2記載において、前記音声符号判定部から出力される前記標準音声パターンと前記声紋とを合成し、さらに前記音の強弱および高低を付して合成音を作成する音声合成部と、この音声合成部によって作成された合成音を記憶する音声記憶部と、この音声記憶部に記憶される合成音を音声信号に変換して前記音声出力手段から出力する音声変換出力部と、前記音声記憶部に記憶される合成音を読み出して前記音声出力手段から繰り返し出力させる音声繰り返しスイッチと、前記音声変換出力部から出力される音声信号の速度および強弱の何れか一方または両方を可変する音声可変手段とを付加したことを特徴とする携帯用音声認識出力補助装置。

【請求項5】 音声入力手段および音声出力手段を有する音声入出力装置部分と、前記音声入力手段から入力された振動周波数信号から声紋、音の強弱および高低、発生音を認識する音声認識部、この音声認識部によって認識された発生音の音声符号と予め記憶されている複数の音声符号とを比較し、両音声符号が一致したとき当該音声符号に対応する予め記

憶されている標準音声パターンを読み出し、この標準音声パターン、前記声紋、音の強弱および高低等の音声情報を発生する音声符号判定部、これら標準音声パターン、声紋、音の強弱および高低等を合成する音声合成部、この音声合成部によって作成された合成音を音声信号に変換し前記音声出力手段から出力する音声変換出力部をもつ本体装置部分と、

前記音声記憶部に記憶される合成音を読み出して前記音声出力手段から繰り返し出力させる音声繰り返しスイッチ、前記音声変換出力部から出力される音声信号の速度および強弱の何れか一方または両方を可変する音声可変手段をもつ音声調整部分とに分けることを特徴とする携帯用音声認識出力補助装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、例えば戸籍簿持主や身体障害者等のことと不明瞭な音声を発声する者が利用して好適な携帯用音声認識出力補助装置に係わり、特に手軽に実行可能とし、また不明瞭な音声を適切に認識して会話の補助に役立てうる携帯用音声認識出力補助装置に関する。

【0002】

【従来の技術】 従来の音声認識装置は、予め発声者の発声する音声に対応する多数の標準音声パターンを記憶し、その発声者の口から発声する音声パターンと予め記憶される多数の標準音声パターンとを比較照合し、発声者の発声する音声パターンと一致する標準音声パターンがあれば、当該標準音声パターンから発声者の発声する音声を認識することが行われている。

【0003】 一方、音声合成装置は、音声を出力する装置であって、アナウンスが発声する音声を録音し、それを分析手法によって低ビットに圧縮、記録し、さらに出力するときに再生する方式と、入力する仮名に対応して単語を組み合わせ、アクセントとイントネーションとを調整する規則合成方式とがある。前者は音声応答装置の出力として利用され、プッシュボンのPB入力和組合せてオーダエントリ分野で利用されている。後者は、日本語、英語の文章から直接音声に変換する技術が開発されており、今後の技術の発展に期待するところが大きい。

【0004】

【発明が解決しようとする課題】 しかしながら、従来の音声認識装置は、発声者の口から発声する音声パターンから音声認識を行っており、例えば手術等で声帯を抽出した者、舌ガンにより舌をなくした者、不明瞭な音声を発声する非発声者等から発する音声については全く認識できない。その理由は、不明瞭な音声を発声するために認識不可能となるだけでなく、口から発声する音声の空気振動を検出しているため、声帯を抽出した者や舌ガンで舌をなくした者の場合にはもともと口から音声を発声

(3)

石嘴平7-13582

9

しないので適用不能となる。

【0005】なお、今後の技術的進歩いかによつては不定多数の音声認識が可能となつたり、また音声認識装置を用いた種々の装置が日常生活の中で使用されてくるであろうが、何れにせよ、健常者に有効な装置の開発であると考えられる。ゆゑに、種々の障害をもつ非健常者は、その音声が不明瞭であつたり、音声の発生速度が遅いために、斬角新しい装置が開発されてもそれを充分に使いこなすことは非常に難しいと思われる。

【0006】一方、前記音声合成装置の場合には、個人の発声する多くの言葉や感情のこもった音声信号とはならず、会話するという観点からみれば余り不十分なものである。

【0007】本発明は上記実情に鑑みてなされたもので、口から音声を発音できない者でも音声に相当する信号を確実に入力可能な携帯用音声認識出力補助装置を提供することを目的とする。

【0008】また、本発明の他の目的は、非健康者の発声する不明瞭な音声を正しく認識し、音声を発声する者の感情を含めた音声合成を実現する携帯用音声認識出力補助装置を提供することにある。

【0009】さらに、本発明の他の目的は、非健常者の身体の状態を考慮しつつ適切な音声を発生する携帯用音声認識出力補助装置を提供することにある。さらに、本発明の他の目的は、非健常者が手軽に装置でき、また操作性に富んだ携帯用音声認識出力補助装置を提供することにある。

[0 0 1 0]

【課題を解決するための手段】上記課題を解決するために、請求項１に対応する発明は、振動発生体に巻付け固定する吸着性の布地で形成された矩冊状の基音体の裏面側に当該振動発生体から発生する振動を検出して電気的な振動周波数信号に変換する平坦状の音声入力手段を取り付け、さらに前記基音体の裏面側に前記振動周波数信号に応じた音声信号を出力する平坦状の音声出力手段を取り付けた音声入出力装置を有する携帯型音声認識出力補助装置である。

【0011】次に、請求項2に対応する発明は、振動発生体から発生する振動を検出して電気的な振動周波数信号を出力する音声入力手段およびこの音声入力手段によって入力された振動周波数信号に応じた音声信号を出力する音声出力手段とを有する音声入出力装置と、前記音声入力手段から入力された振動周波数信号から声紋、音の強弱および高低、発生音を認識する音声認識部と、予め複数の標準音声パターンおよび当該パターンに対応する音声符号が記憶され、前記音声認識部によって認識された発生音に係わる音声符号と較べて記憶されている前記音声符号とを比較し、両音声符号が一致したとき、前記音声符号に対応する標準音声パターンを読み出し、当該標準音声パターン、前記声紋、音の強弱および高低等が

らなる音声情報を入力する音声符号判定部とを設けた携帯用音声認識出力補助装置である。

【0012】次に、請求項3に対応する発明は、請求項2に対応する発明の構成要件に、新たに前記音声符号判定部から出力される前記標準音声パターンと前記声紋とを合成し、さらに前記音の強弱および高低を付して合成音を作成する音声合成部と、この音声合成部で作成された合成音を音声信号に変換して前記音声出力手段から出力する音声変換出力部とを付加してなる携帯用音声認識出力補助装置である。

【0013】さらに、請求項1に対応する発明は、請求項2に対応する発明の構成要件に、新たに前記音声符号判定部から出力される前記標準音戸パターンと前記戸数とを合成し、さらに前記戸の強弱および高低を付して合成音を作成する音声合成部と、この音声合成部によって作成された合成音を記憶する音声記憶部と、この音声記憶部に記憶される合成音を音声信号に変換して前記音声出力手段から出力する音声変換出力部と、前記音声記憶部に記憶される合成音を読み出して前記音声出力手段から繰り返し出力させる音戸繰返しスイッチと、前記音声変換出力部から出力される音声信号の歪度を可変する音声歪度可変手段と、前記音声変換出力部から出力される音声信号レベルを可変し強調を付ける音戸強調可変手段とを付加してなる標準音戸変換出力補助装置である。

【0014】さらに、請求項5に対応する発明は、音戸人力手段および音声出力手段とを有する音声入力装置部分と、音戸認識部、音声符号判定部、音声変換出力部をもつ本体装置部分と、前記音戸認識部に記憶される合成音を読み出して前記音声出力手段から繰り返し出力させる音声繰り返しスイッチ、前記音声変換出力部から出力される音声信号の速度を変変する音声速度可変手段、前記音声変換出力部から出力される音声信号レベルを変変し強調を付ける音戸強調可変手段をもつ音声調整部分とに分けた携帯用音戸認識出力補助装置である。

[0015]

【作用】従つて、請求項1に対応する発明は以上のような手段を調じたことにより、振動発生体、例えば非随常者の首に巻き付け固定する装着体に吸音性の布地を用い、かつ、装着体の裏面側および前面側とにそれぞれ個別に平坦状の音声入力手段および音声出力手段を取り付けたことにより、口から発声する音声や外部から入ってくる雑音の影響を防止でき、しかも非随常者の喉に対する負担が軽減され、喉付近から発声する振動を確実に入力することができ、

【0016】次に、請求項2に対応する発明は、音声認識部が音声入力手段から入力される振動周波数成分から声紋、音の強弱、音の高低および発声位置を認識して音声符号判定部へ送出する。この音声符号判定部では、予め演算の標準音声パターンおよび当該パターンに対応する音声符号が記憶されているので、音声認識部から送られ

(4)

特快早7-13582

5

てくる発声音に係わる音符号と既に記録されている音符号とを比較し、両音符号が一致したとき、その音符号に対応する標準音声パターンを読み出し、当該標準音パターン、前記音紋、音の強弱および高低等からなる音声情報を読み出すので、非標準音の発声する不明瞭な音でも正しく認識でき、また非標準音の発声する短い言葉から日常会話等に用いる長い言葉に要換されている標準音パターンを容易に出力できる。

【0017】さらに、請求項3に対応する発明は、請求項2に対応する発明と同様な作用を有する他、音声合成部にて音声符号判定部から送られてくる標準音声パターンと前記声紋とを合成し、さらに音の強弱、高低を付して合成音を作成するので、感情を含めて音声合成でき、しかも音声信号変換出力部において合成音を音声信号に変換して前記音声出力手段から出力するので、感情表現を伴った音声信号を出力できる。

【0018】さらに、請求項4に対応する発明は、請求項2および請求項3に対応する発明と同様な作用を有する他、音戸繰返しスイッチを操作して前記音声記憶部から再度合成音を読み出して音声出力手段から繰り返し出力するので、相手から聞き直された場合でも最初から音声を発することなく同様の音声信号を出力できる。また、音戸速度可変手段によって音声信号の出力速度を可変することにより、健常者にとって分かり易い速度で音声信号を出力できる。また、音戸強弱可変手段によって音声信号レベルを可変し強弱を付けて出力するので、同様に聴覚者にとって分かり易い音声信号を出力できる。

【0019】さらに、請求項5に対応する発明は、音声入力手段および音声出力手段とを有する音声入出力装置部分と、音声認識部、音声符号判定部、音声信号変換出力部等をもつ本体装置部分と、種々の調整機能をもつ音声調整部分とに分けることにより、音声入出力装置部分は非異常者の首に巻き付け、本体装置部分は胴体の腰部などに吊下し、音声調整部分は手元に掲げて操作できるようにすれば、簡単に操作でき、かつ、手元に操作できる。

[0 0 2 0]

【実施例】以下、本発明の実施例について図面を参照して説明する。図１は本発明装置の構成を示すブロック図である。図１において１は音声入出力装置であって、これは図２に示すごとく例えばむち打ち症などのときに首に巻き付けるコルセットのような例えば布地の装着体１１が用いられ、この装着体１１の適宜な箇所には喉から発する振動を直接取り込む音声入力手段１２および音声信号を出力する音声出力手段１３が取り付けられ、さらに首に巻き付け固定するために装着体両端部の対向面にマジックテープ１４a、１４bが取り付けられている。なお、マジックテープ１４a、１４b以外の従来周知の種々の固定手段例えばホックなどを用いて固定してもよい。

6

【0.0:2.1】拘記装着体1.1は、例えば外部雑音を遮断するカーテン等のごとく吸音性の優れた布地で作成し、これによって口から発声する音声や外部から入ってくる雑音を吸収し、前記音声入力手段1.2に影響を与えないようにする。音声入力手段1.2は、装着体1.1の鼻咽（内側）面部に平坦に取り付けられ、喉から発声する振動を電気信号に変換して出力する。このように平坦化することにより装着体1.1に馴染み易く、喉への圧迫感がなく、ひいては喉に対する負担を軽減できる。一方、音声出力手段1.3は、音声入力手段1.2とは反対側、つまり装着体1.1の鼻咽（外側）面部に同様にフラットなスピーカを取り付けられる。このようにフラットなスピーカを口と同じ縦ライン上の正面に取り付けることにより、喉に対する負担が軽減され、話し相手からみればあたかも口から音声が発する状態を作り出す。また、この音声出力手段1.3は、装着体1.1と同系色または適宜な素材で覆うとか、開口出力手段1.3の色に適宜な工夫を講じることにより、目立つ限り目立たない自然な取り付け状態に取り付けられるものとする。

20 【0022】2は音声入力手段12から入力される音声振動周波数信号から個人の声紋の特徴、音の強弱と高低、正しい発声を認識する音声認識部である。この音声認識部2は、図3に示すように音声スペクトル変換手段21、音質判定手段22、声紋判定手段23および発声音高検出手段24等からなっている。この音声スペクトル変換手段21は、例えば図4(a)に示すような音声振動周波数成分を所定の周期でサンプリングすることにより、図4(b)に示すような音声スペクトルに変換する。音質判定手段22は、音声スペクトルから音の強弱と高低とを判定するものであり、そのうち音の強弱は、予め所定の基準レベルが設定され、音声スペクトルの各成分が基準レベルから上下方向にどの程度レベル的に離れているかを表すものであり、一方、音の高価は音の周波数に依存するが、ここでは専ら音声スペクトルの各成分のレベルを表す。声紋判定手段23は音声スペクトルの周波数成分レベルを抽出するものであり、また発声音認識手段24は音声スペクトルの分布状態から発声音を決定し、その発声音に対応する文字コード、例えば「ア」とか「イ」とかのコードに変換し出力する。そして、これら判定手段22~24によって判定されたデータは時系列的に出力され、音声符号判定部3に送られる。

【0023】この音声符号判定部3は、予め標準音声パターンとそれに対応する音声符号とが記憶され、発声音認識手段24にて音声認識された正しい発声音である文字コード（音声符号）を取り出し、この音声符号と既に記憶されている音声符号とを比較し、両音声符号が同一となるとき、それに対応する標準音声パターンを出力する機能を有する。具体的には、図5に示すように標準音声パターンを記憶する音声パターン記憶手段31と、

(5)

特開平7-13582

7

この音声パターン記憶手段31の各標準音声パターンに対応する音声符号を記憶する音声符号記憶手段32と、音声符号判定手段33とによって構成されている。

【0024】この音声符号判定手段33は、前記音質判定手段22からの音の強弱、高低に関するデータおよび声紋判定手段33からの声紋の特徴データをバッファメモリ待ちの状態にし、発声音認識手段24で認識された正しい発声音の音声符号については、当該音声符号と音声符号記憶手段32に記憶されている多数の音声符号とを比較参照し、既に記憶されている音声符号と同一であれば、音声パターン記憶手段31から音声符号に対応する標準音声パターンを取り出し、既にバッファメモリ待ちの状態にあるデータとともに音声情報記憶部4に記憶する。このとき、発声音認識手段24の発声音の音声符号も同時に記憶してもよい。一方、発声音認識手段24によって認識された音声符号と既に記憶されている音声符号とが不一致となったとき、その発声音認識手段24で認識された発声音の音声符号を出力する。

【0025】なお、前記音声パターン記憶手段31に記憶されている標準音声パターンは、例えば“おはようございます”、“ありがとうございます”、“さようなら”などの日常会話で使用する言葉に相当するパターンである。つまり、短い音声符号から長い言葉に変換することにより、非健常者が全ての言葉を発声しなくても十分に会話可能にパターン化している。

【0026】前記音声情報記憶部4は、声紋の特徴、音の強弱、音の高低および発声音に係わる標準音声パターン、必要に応じて認識された発声音の音声符号などの音声情報を一時的記憶した後、音声合成部5に送出する。

【0027】この音声合成部5においては、図6に示すように音声情報記憶部4から送られてくる音声情報を記憶する音声情報記憶手段51と、この音声情報記憶手段51に記憶されている音声情報のうち、標準音声パターンと声紋の特徴データとを合成し、さらにかかる合成音に音の強弱および音の高低を付けることにより、完全に擬似化した合成音を作り出し、後述の音声記憶部6に記憶する音声合成手段52とで構成されている。

【0028】7は音声変換出力部であって、これは音声記憶部6に記憶されている合成音情報を読み出して音声出力可能なアナログ信号に変換して音声出力手段13から音声を出力する機能をもっている。

【0029】さらに、本装置には音声出力調整部8が設けられている。この音声出力調整部8を設けた理由は、非健常者の状況に応じて会話の内容が相手側に適切に伝達できるようにすることにある。すなわち、音声出力調整部8には、一度、音声出力手段13から出力された音声信号が相手側から聞き直されたとき、音声記憶部6から繰り返し合成音を出力させるために読み出し操作を行う音声繰り返しスイッチ81が設けられている。これは、非健常者が最初から同じ音声を発声するのが非常に大変

9

であるので、その負担を軽減するためである。

【0030】また、この音声出力調整部8には、音声速度可変器82および音声強弱可変器83が設けられている。予め音声変換出力部7側にコンデンサなどを付いたアナログ的な1次遅れ回路を組み込んでおき、音声速度可変器82で適宜に1次遅れ回路を短絡することにより、音声信号の速度を変換する。これは非健常者の発声速度は必ずしも早くないので、音声出力手段13から出力される合成音の出力速度を適宜変更し、健常者が聞き取り易い速度にするためである。また、音声強弱可変器83は、音声変換出力部7側の音声信号のレベルを変換するとか、増幅率を変換することにより、音声信号に強弱を付けて出力する。これは外部の雑音が多いところでも音声出力手段13から出力される音声信号に強弱を付けて聞き取り易くするためである。

【0031】次に、以上のように構成された装置の動作について説明する。まず、非健常者が音声入出力装置1の装着体11を首に巻き付けた後、装着体11の両端対峙面に設けたマジックテープ部分を押し付けて固定する。このとき、装着体11に取り付けられている音声出力手段13が正面位置にくるように設定し、また音声入力手段12は喉の振動を最も取り込み易い部位、例えば首の側部の位置に設定する。このとき、音声入力手段12および出力手段13が平直状に形成されているので、首に馴染み易く、喉に対する負担が非常に少なくなる。

【0032】この状態において非健常者が音声を発生すると、当該非健常者の喉の振動を音声入力手段12で取り込んで電気的な振動周波数信号に変換し、音声認識部2に送出する。

【0033】ここで、音声認識部2は、音声入力手段12から入力される振動周波数信号を音声スペクトル変換手段21により音声スペクトルに変換した後、音質判定手段22、声紋判定手段23および発声音判定手段24に送出する。これら各判定手段22～24は前述した判定条件に従って音の強弱および音の高低、声紋の特徴および正しい発声音を決定し、特に発声音の場合には発声音に対応する文字コード（音声符号）に変換し、音の強弱および音の高低、声紋の特徴データとともに音声符号判定部3に送出する。

【0034】この符号判定部3においては、予め音声パターン記憶手段31に標準音声パターンが記憶され、また音声符号記憶手段32に前記標準音声パターンに対応する音声符号が記憶されており、特に標準音声パターンには例えば“おはようございます”、“ありがとうございます”、“さようなら”などの日常会話で使用する言葉に相当するパターンの形で保存されている。

【0035】従って、符号判定部3では、音声認識部2によって認識された正しい発声音である文字コード（音声符号）を受けると、そのいくつかの音声符号と既に記憶されている音声符号とを比較し、両音声符号が同一とな

(6)

特開平7-13582

9

ったとき、それに対応する標準音声パターンを読み出し、前記音質判定手段22からの音の強弱、高低に関するデータおよび声紋判定手段23からの声紋の特徴データとともに音声情報記憶部4を介して音声合成部5に送出する。

【0036】ここで、音声合成部5は、音声情報記憶部4から送られてくる標準音声パターン、音の強弱、高低および声紋等の音声情報を音声情報記憶手段51に一旦記憶した後、音声合成手段62で音声合成を行う。この音声合成は、音声情報のうち、標準音声パターンと声紋の特徴データとを合成し、さらにかかる合成音に音の強弱および音の高低を付けて完全な復調をなした合成音を作り出し、音声記憶部6に記憶した後、音声変換出力部7に送られる。この音声変換出力部7では、音声記憶部6に記憶されている合成音信号を読み出して音声出力可能なアナログ信号に変換して音声出力手段13から音声

【0037】このとき、例えば相手側から聞かされたとき、非健常者は、音声送返しスイッチ81を操作すれば、音声記憶部6から再度合成音情報を読み出し、音声変換出力部7にて音声出力可能なアナログ信号に変換して音声出力手段13から音声を出力するので、相手側に適切な音声信号、つまり会話の内容を伝えることができる。また、非健常者の発声速度が遅い場合には、音声速度可変器82で適宜に音声信号の出力速度を早くすれば、健常者等が聞き取り易くなる。また、例えば外部の雑音が多いところでは、音声強調可変器83を可変操作すれば、音声信号レベルを大きくして音声出力手段13から出力でき、同様に健常者等が聞き取り易くなる。

【0038】従って、以上のような実施例の構成によれば、音声入出力装置1の本体となるべき装着体11は吸着性に優れた布地などで作成したので、非健常者の首に巻き付けたときに完全になじむだけでなく、口から発声する音声や外部から入ってくる雑音を吸収し、音声入力手段12からは喉から発声する振動を適切に入力できる。しかも、装着体11の面部には平坦状の音声入力手段12および音声出力手段13を貼り付けるように取り付ければ、軽便で携帯に便利であり、喉に対する圧迫感などがなくなり、喉に対する負担を軽減できる。また、音声認識部2において音声入力手段12から入力される振動周波数信号から声紋の特徴、音の強弱および音の高低、発声音を認識し、この発声音の音声符号と声紋の特徴、音の強弱および音の高低情報等を音声符号判定部3に送出し、ここで音声符号と予め記憶されている多数の音声符号とを比較し、両音声符号が一致するとき、当該音声符号に対応するありがとうございます、さようならなどの日常会話で使用する言葉に相当する標準音声パターンを読み出し、前記声紋の特徴、音の強弱および音の高低等とともに音声合成部5に送出するようにしたので、非健常者による最初の短い会話の発声から

10

日常会話である長文の標準音声パターンを出力でき、非健常者による会話の負担を十分に補助できる。

【0039】さらに、音声合成部5において、音声符号判定部3側から送られてくる各種の音声情報を一旦記憶した後、その音声情報の中から標準音声パターンに声紋の特徴を合成し、さらに音の強弱および音の高低を付けたので、非健常者の感情を含めた合成音を作成できる。

【0040】さらに、音声信号を送り返し出力する音声送返しスイッチ81、音声信号の速度や強度を可変する音声速度可変器82や音声強調可変器83を設けたので、非健常者の状況や相手側の聞き取り状態に応じて適宜に操作しながら適切な音声信号を出力できる。

【0041】なお、上記実施例では、全体の構成について述べたが、非健常者が手軽に携帯し簡単に操作する観点から考えたとき、次のような分割構成とすることが望ましい。つまり、音声入力手段12および音声出力手段13を有する音声入出力装置部分と、音声認識部2、音声符号判定部3、音声情報記憶部4、音声合成部5、音声記憶部6および音声変換出力部7等からなる電源部分を含む装置本体部分と、音声送返しスイッチ81、音声速度可変器82および音声強調可変器83等の音声出力調整部分とに分割すれば、適宜に信号線で接続するようにすれば、音声入出力装置部分を首に巻き付け固定し、装置本体部分を腰に吊下し、音声出力調整部分を手にもつことができ、これによって手軽に携帯でき、操作性を上げることができる。

【0042】また、装着体11は、布地を用いたが、吸着性の紙地またはそれに類する素材であれば、特に限定するものではない。その他、本発明はその要旨を逸脱しない範囲で種々変形して実施できる。

【0043】

【発明の効果】以上説明したように本発明によれば、次のような種々の効果を奏する。請求項1の発明においては、口から音声を発声できない者でも音声に相当する信号を確実に入力でき、かつ、非健常者の喉を圧迫せずに喉の振動を適切に入力できる。

【0044】請求項2、3の発明は、非健常者の発声する不明瞭な音声を正しく認識でき、しかも音声パターン、声紋および音の強弱等を合成することにより、音声を発声する者の感情を含めた音声合成を実現できる。

【0045】次に、請求項4の発明は、非健常者の身体の状態を考慮し、かつ、相手の聞き取り状態に応じて適宜に音声操作を行って適正な音声信号を発生することができる。さらに、請求項5の発明は、構成を適切に分割することにより、非健常者が手軽に装着でき、また非健常者による操作性を高めることができる。

【図面の簡単な説明】

【図1】本発明に係わる携帯用音声認識出力補助装置の一例を示す機能ブロック図。

【図2】図1に示す音声入出力装置の構成を示す図。



(7)

特開平7-13582

11

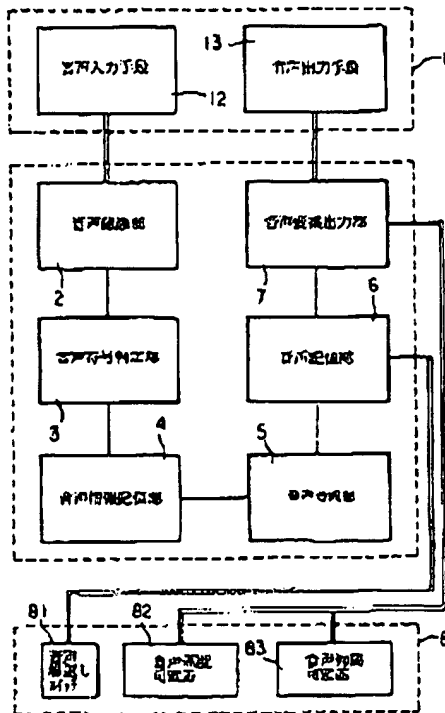
【図3】図1に示す音声認識部を具体化した機能ブロック図。

【図4】音声認識部による音声認識を説明する図。

【図5】図1に示す音声符号判定部を具体化した機能ブロック図。

【図6】図1に示す音声合成部を具体化した機能ブロック図。

【図1】

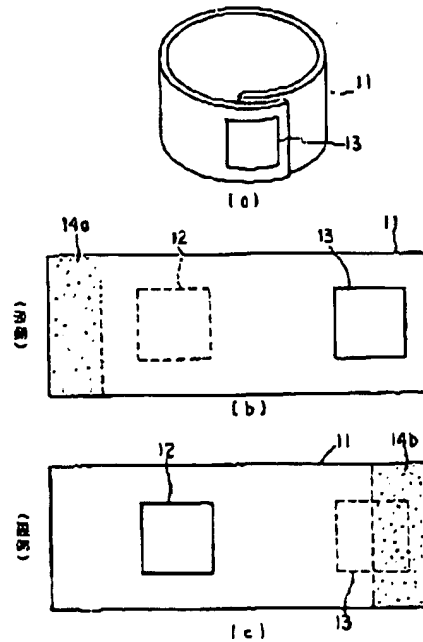


12

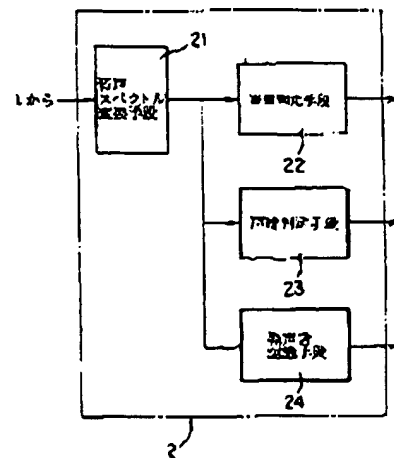
【符号の説明】

1…音声入出力装置、2…音声認識部、3…音声符号判定部、4…音声情報記憶部、5…音声合成部、6…音声記憶部、7…音声変換出力部、8…音声出力調整部、11…装着体、12…音声入力手段、13…音声出力手段、81…音声認識部、82…音声認識部、83…音声認識部。

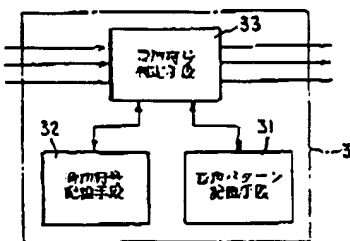
【図2】



【図3】



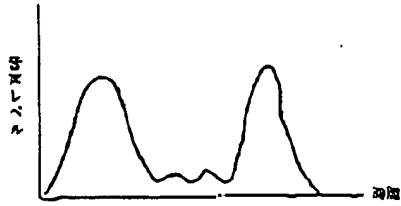
【図5】



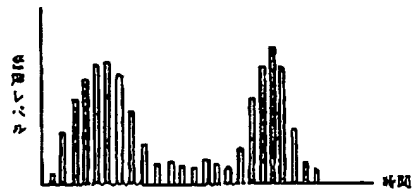
(8)

符号率 7-13582

【図4】



(a)



(b)

【図6】

